

IOWA STATE UNIVERSITY

Digital Repository

Economics Working Papers (2002–2016)

Economics

4-14-2008

An axiomatization of the multigroup Atkinson segregation indices

David M. Frankel

Iowa State University, dfrankel@iastate.edu

Oscar Volij

Iowa State University; Ben Gurion University, ovolij@bcu.ac.il

Follow this and additional works at: http://lib.dr.iastate.edu/econ_las_workingpapers



Part of the [Economics Commons](#)

Recommended Citation

Frankel, David M. and Volij, Oscar, "An axiomatization of the multigroup Atkinson segregation indices" (2008). *Economics Working Papers (2002–2016)*. 165.

http://lib.dr.iastate.edu/econ_las_workingpapers/165

This Working Paper is brought to you for free and open access by the Economics at Iowa State University Digital Repository. It has been accepted for inclusion in Economics Working Papers (2002–2016) by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

An axiomatization of the multigroup Atkinson segregation indices

Abstract

This paper gives an axiomatic characterization of the multigroup Atkinson indices of segregation relying entirely on ordinal axioms. We show that the Symmetric Atkinson index represents the unique ordering that treats ethnic groups symmetrically, that is invariant to population growth rates that differ among ethnic groups, that regards school districts as more segregated when schools in them are subdivided (unless the new schools have the exact same ethnic distribution), and that satisfy an independence property. If symmetry among ethnic groups is dropped, one obtains the family of orderings that are represented by the Asymmetric Atkinson indices. The latter result requires the addition of a continuity axiom.

Keywords

education, segregation, schools, diversity

Disciplines

Economics

IOWA STATE UNIVERSITY

An Axiomatization of the Multigroup Atkinson Segregation Indices

David M. Frankel, Oscar Volij

April 2008

Working Paper # 08010

Department of Economics Working Papers Series

Ames, Iowa 50011

Iowa State University does not discriminate on the basis of race, color, age, religion, national origin, sexual orientation, gender identity, sex, marital status, disability, or status as a U.S. veteran. Inquiries can be directed to the Director of Equal Opportunity and Diversity, 3680 Beardshear Hall, (515) 294-7612.

An Axiomatization of the Atkinson Segregation Indices*

David M. Frankel

Oscar Volij

Iowa State University

Iowa State University

and Ben Gurion University

April 14, 2008

Abstract

This paper gives an axiomatic characterization of the Atkinson indices of segregation for the multigroup case using a small number of purely ordinal axioms. We show that the symmetric Atkinson index represents the unique ordering that treats ethnic groups symmetrically, that is invariant to population growth rates that differ among ethnic groups, that ranks school districts as more segregated when schools in them are subdivided (unless the new schools have the exact same ethnic distribution), and that satisfy an independence property. If symmetry among ethnic groups is dropped and a technical continuity axiom added, one obtains the family of orderings that are represented by the asymmetric Atkinson indices.

*Email addresses: dfrankel@econ.iastate.edu; ovolij@bgu.ac.il. Volij thanks the Spanish Ministerio de Educación y Ciencia (project SEJ2006-05455) for research support.

1 Introduction

Empirical research indicates that segregation affects economic outcomes. For instance, Cutler and Glaeser [6] find that residential racial segregation leads to higher dropout, idleness, and single motherhood rates among African Americans. Others have found that occupational segregation by gender helps explain the gender gap in wages and that racial segregation of schools helps explain the black-white achievement gap.¹

While there seems to be a consensus that segregation is important, there is less agreement about how to define it. Massey and Denton [19] discern five dimensions of segregation. The first, evenness, is the tendency of ethnic groups to be distributed differently across locations, such as neighborhoods or schools. This is also the definition favored by James and Taeuber [15]. Massey and Denton's other dimensions are isolation from the majority group, concentration in a small area, centralization in the urban core, and clustering in a contiguous enclave.

In this paper, we will focus exclusively on Massey and Denton's first dimension of evenness. How best to measure this dimension of segregation remains an open question. A number of indices have been proposed (Massey and Denton [19]). Selected properties of these indices have been studied.² However, this leaves some questions unanswered. Can other indices be devised that also satisfy these properties? And what other properties do the existing indices satisfy?

A more definitive approach is to provide an axiomatization: a set of properties A that are satisfied by all and only the indices in some set. This answers the above two questions: no other indices satisfy all of properties in A ; any other properties of the indices in the set must be implied by the properties in A .

This paper provides two axiomatizations: one for the symmetric Atkinson index, and one

¹For gender segregation, see Cotter et al [5], Lewis [17], and Macpherson and Hirsh [18]. For school segregation, see Boozer, Krueger, and Wolkon [2] and Hanushek, Kain, and Rivkin [11].

²See, for instance, Duncan and Duncan [8], James and Taeuber [15], Massey and Denton [19], Reardon and Firebaugh [21] for the multigroup case, and Zoloth [25].

for the set of asymmetric Atkinson indices. The Atkinson indices were introduced by James and Taeuber [15] and are based on the Atkinson family of inequality indices (Atkinson [1]). Massey and Denton [19] study properties of the Atkinson indices and Johnston, Poulsen, and Forrest [16] use them to study school segregation. While this literature has focused on the case of two ethnic groups, we study the general multigroup case.

In this paper we will also focus on contexts in which geography is unimportant. In some cases, such as residential neighborhoods, this might be a strong assumption. In others, it is more innocuous. For instance, the presence of other schools near a given student's school typically does not have a great effect on the student's educational outcomes. Hence, our presentation will focus on school district segregation. However, it should be clear that the measures of segregation characterized here could also be applied to measure segregation in other contexts such as gender segregation in the labor force.

We also restrict to ordinal axioms. Ordinal axioms are more appealing than cardinal ones because they refer to bilateral comparisons and not to their specific functional representations. Formally, we define a segregation ordering as a ranking of school districts from least segregated to most segregated. We show that the ordering that is captured by the symmetric Atkinson index is the unique nontrivial ordering that satisfies the following axioms: Group Symmetry, Scale Invariance, the Weak School Division Property, and Independence.

Informally, these axioms are defined as follows. Symmetry requires that the segregation ordering be invariant to the renaming of the groups. Scale Invariance states that the segregation ranking of a school district should depend only on how the different ethnic groups are distributed across schools, and not on the absolute sizes of these groups. The Weak School Division Property states that in a school district that contains a single school, building a new school to which some of the students are moved

1. cannot lower segregation in the district, and
2. leaves segregation unchanged if the ethnic distributions of the two resulting schools are identical.

Lastly, Independence states that the segregation ranking of two school districts with the same size and ethnic distribution is unaffected by the addition of one identical school to each of the two districts.

A simple representation of the symmetric Atkinson ordering is the Hutchen's [13] Square Root Index: one minus the sum, over all schools, of the geometric averages of the percentages of each group who attend the school. For instance, in the case of two ethnic groups, suppose 40% of blacks and 10% of whites attend school A while 60% of blacks and 90% of whites attend school B. The index equals $1 - (.4)^{1/2} (.1)^{1/2} - (.6)^{1/2} (.9)^{1/2} = 0.065$. Abusing terminology, we will call this the symmetric Atkinson index.³

We then drop Symmetry and show that all and only the asymmetric Atkinson orderings satisfy the remaining axioms, with the addition of a technical Continuity axiom. Each asymmetric Atkinson ordering is represented by the following index: one minus the sum, over all the schools, of some *weighted* geometric average of the percentages of each group who attend the school. In the above example this would equal $1 - (.4)^b (.1)^{1-b} - (.6)^b (.9)^{1-b}$ where $b \in (0, 1)$ is the weight given to blacks. The parameters b trace out the full family of asymmetric Atkinson orderings in the case of two ethnic groups.⁴

The paper is organized as follows. Concepts and notation are defined in Section 2. Section 3 gives examples of segregation indices. Section 4 presents the axioms. Results appear in section 5. In section 6, we conclude and discuss related literature. Proofs are relegated to an appendix.

³The original Atkinson index is an increasing transformation of this index and thus captures the same ordering (section 3).

⁴Throughout, we use "asymmetric" as shorthand for "not necessarily symmetric." In particular, this set includes the symmetric Atkinson index ($b = 1/2$).

2 Definitions

We assume a continuum population. This technical assumption allows us to fully characterize the Atkinson orderings by means of a few axioms. With a discrete population, the Atkinson measures still satisfy our axioms; however, there may be other orderings that do so as well. In practice, this flexibility allows a researcher to assign a different weight to different groups of people; for instance, in the case of residential segregation, one might want to assign different weights to children vs. adults.

Formally, we define a (school) district as follows:

Definition 1 A *district* consists of

- A finite set of groups G , containing at least two elements,
- a nonempty and finite set of schools N ,
- and, for each ethnic group $g \in G$ and for each school $n \in N$, a nonnegative number $T_g^n \in \mathfrak{R}_+$, representing the number of members of group g that reside in school n ,⁵

such that the total population of group g in the district is positive: for all $g \in G$, $\sum_{n \in N} T_g^n > 0$.

Fix an integer $K > 1$ and let $\mathcal{C} = \mathcal{C}(K)$ be the set of districts whose set of ethnic groups, G , contains precisely K groups. A *segregation ordering* \succsim on the set of districts \mathcal{C} is a complete and transitive binary relation on that set. We interpret $X \succsim Y$ to mean “district X is at least as segregated as district Y .” The relations \sim and \succ are derived from \succsim in the usual way.

A *segregation index* on \mathcal{C} is a function that assigns a nonnegative number to each district $X \in \mathcal{C}$. Any segregation index S induces a segregation ordering defined by $X \succsim Y \Leftrightarrow$

⁵For instance, $T_g^n = 127$ means that neighborhood n contains 127 members of ethnic group g .

$S(X) \geq S(Y)$. As in utility theory, a segregation ordering may be represented by more than one index, and there are segregation orderings that are not captured by any index.

For any district $X = \langle (T_g^n)_{g \in G} \rangle_{n \in N}$, we denote the set of schools of X by $N(X)$. We will sometimes use a more compact notation. The expression $\langle (1, 2), (3, 1) \rangle$, for instance, denotes a district with two ethnic groups (e.g., blacks and whites) and two schools. The first school, $(1, 2)$, contains one black and two whites; the second, $(3, 1)$, contains three blacks and one white. The order of the schools does not matter; e.g., $\langle (1, 2), (3, 4) \rangle$ can also be written $\langle (3, 4), (1, 2) \rangle$.

The following notation will be useful:

$$\begin{aligned}
T_g &= \sum_{n \in N} T_g^n: \text{ the number of members of group } g \text{ in the district} \\
T^n &= \sum_{g \in G} T_g^n: \text{ the total population of school } n \\
T &= \sum_{g \in G} T_g: \text{ the total population of the district} \\
P_g &= \frac{T_g}{T}: \text{ the proportion of district students who are in group } g \\
P^n &= \frac{T^n}{T}: \text{ the proportion of district students who are in school } n \\
p_g^n &= \frac{T_g^n}{T^n} \text{ (for } T^n > 0\text{): the proportion of students of } n \text{ who are in } g \\
t_g^n &= \frac{T_g^n}{T_g}: \text{ the proportion of members of } g \text{ who attend } n
\end{aligned} \tag{1}$$

The *group distribution of a district* X is the vector $(P_g)_{g \in G}$ of proportions of the district's students who are in each group. The *group distribution of a nonempty school* n is the vector $(p_g^n)_{g \in G}$ of proportions of the school's students who are in each group. School n in district X is *representative* if the group distributions of n and X are the same: if $p_g^n = P_g$ for all $g \in G$. A school that is not representative of the district is said to be *unrepresentative*.

For any two districts X and Y with the same set of groups, $X \uplus Y$ denotes the result of adjoining Y to X . Its schools are the union of the schools in X and Y . Formally, let $X = \langle (T_g^n)_{g \in G} \rangle_{n \in N}$ and $Y = \langle (T_g^n)_{g \in G} \rangle_{n \in N'}$ with disjoint set of neighborhoods. Then $X \uplus Y$ denotes the city $\langle (T_g^n)_{g \in G} \rangle_{n \in N \cup N'}$.

3 Examples of segregation indices

We now discuss several examples of segregation indices. We begin with the symmetric Atkinson index:

Symmetric Atkinson The symmetric Atkinson index A is defined by

$$A(X) = 1 - \sum_{n \in N(X)} \left(\prod_{g \in G} t_g^n \right)^{\frac{1}{|G|}} \quad (2)$$

When X contains exactly two nonempty groups, this index is an increasing transformation of the usual Atkinson index with parameter $1/2$ (Massey and Denton [19, p. 286]).⁶ The symmetric Atkinson index is derived from the income inequality measure of the same name (Atkinson [1]).

Asymmetric Atkinson Let $\mathbf{w} = (w_1 \dots w_K)$ be a vector of K nonnegative weights that sum to one. The asymmetric Atkinson index with weights \mathbf{w} , $A_{\mathbf{w}}$, is defined by

$$A_{\mathbf{w}}(X) = 1 - \sum_{n \in N(X)} \left(\prod_{g \in G} t_g^n \right)^{w_g} \quad (3)$$

The weights may all be equal, in which case $A_{\mathbf{w}}$ is just the symmetric Atkinson index.

Unweighted Dissimilarity The Unweighted Dissimilarity index $D^U : \mathcal{C} \rightarrow [0, 1]$ is defined by

$$D^U(X) = \frac{1}{2(K-1)} \sum_{n \in N(X)} f(t^n) \text{ where } f(t^n) = \sum_{g \in G} \left| t_g^n - \sum_{g' \in G} \frac{1}{K} t_{g'}^n \right| \quad (4)$$

In the case of two groups, this index measures the proportion of either group who would have to change schools in order to attain complete integration. This index was

⁶One can show that with two groups, the usual Atkinson index with parameter $1/2$ equals $1 - (1 - A(X))^2$, which is an increasing transformation of A .

first discussed, for the case of two groups, by Jahn *et al* [14]. It was used by Cutler, Glaeser, and Vigdor [7] to measure the evolution of segregation in American cities.⁷

Mutual Information The *entropy* of any discrete probability distribution $q = (q_1, \dots, q_K)$ (where $\sum_{k=1}^K q_k = 1$) is defined by⁸

$$h(q) = \sum_{k=1}^K q_k \log_2 \left(\frac{1}{q_k} \right).$$

The Mutual Information index equals the entropy of the district’s ethnic distribution minus the average entropy of the ethnic distributions of its schools:

$$M(X) = h(P) - \sum_{n \in N} P^n h(p^n)$$

where $P = (P_g)_{g \in G}$ is the district ethnic distribution and $p^n = (p_g^n)_{g \in G}$ is the ethnic distribution of school n . This index is axiomatized in Frankel and Volij [10].

4 Axioms

We impose axioms not on the segregation index but on the underlying segregation ordering. Our first axiom, Group Symmetry, states that the level of segregation in a district does not depend on the labeling of the district’s demographic groups; it depends only on the number of people in each group who attend each school. For instance, if “blacks” are relabeled “whites” and vice-versa, then segregation does not change.

Group Symmetry (GS) Let $X \in \mathcal{C}$ be a district and let $X' \in \mathcal{C}$ be the district that results from relabeling some or all of the groups in X . Then $X \sim X'$.

⁷The two-group Dissimilarity index of Jahn *et al* [14] can be generalized to the multigroup case in various ways. The version in (4) gives equal weight to all groups and satisfies Scale Invariance. An alternative version, which weights a group according to its relative size, is discussed by Reardon and Firebaugh [21] and Frankel and Volij [10]. This alternative version does not satisfy Scale Invariance but does satisfy a different axiom, the Group Division Property, which D^U violates.

⁸When $q_k = 0$, the term $q_k \log_2(1/q_k)$ is assigned the value zero.

We will consider axiomatizations both with and without this axiom.

Arguably, in order to best capture Massey and Denton’s dimension of evenness, a measure would rank a district based solely on how the ethnic groups are distributed across the district’s schools. The measure should not be “contaminated” by differences in the ethnic size distribution from one district to another. This is formalized as follows:

Scale Invariance (SI) For any district $X \in \mathcal{C}$, group $g \in G(X)$, and constant $\alpha > 0$, let X' be the result of multiplying the number of group- g students in each school n in district X by α . Then $X' \sim X$.

Scale Invariance is one of the five requirements that Jahn *et al* [14] say a satisfactory measure of segregation should satisfy.⁹ In their justification of Scale Invariance, James and Taeuber write:

School segregation refers to racial variation in the distribution of students across schools. ... This concept of segregation does not depend on the relative proportions of blacks and whites in the system, but only upon the relative distributions of students among schools.... [Taeuber and James [24, p. 134]]

Scale Invariance can be useful for longitudinal comparisons as it ensures that simple population growth will not affect a district’s segregation ranking. For instance, if the number of blacks grows by 10% in all schools, while each ethnic group retains its distribution across schools, then by this axiom segregation in the district is unaffected.

On the other hand, Scale Invariance may not be suitable in all contexts. For instance, researchers who conceive of segregation as isolation from the majority group (Massey and Denton’s second dimension) have generally not assumed this property (Coleman, Hoffer, and Kilgore [4, p. 178]). For a different axiomatization that does not assume Scale Invariance, see Frankel and Volij [10].

⁹They write: “a satisfactory measure of ecological segregation should ... not be distorted by the size of the total population, the proportion of Negroes, or the area of a city...” (Jahn *et al* [14]).

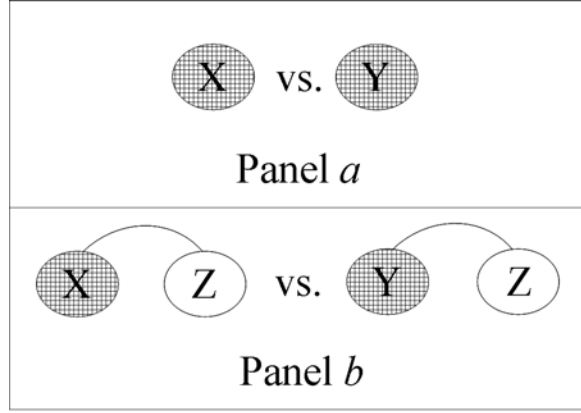


Figure 1: Independence (IND). Panel *a* shows two districts, X and Y , that have the same size and ethnic distribution. IND states that adjoining the same cluster containing a single school to the two districts (panel *b*) does not affect which district is more segregated.

Our next axiom is illustrated in Figure 1. In panel *a*, two districts, X and Y , are being compared. The districts are assumed to have the same number of students from each ethnic group.¹⁰ In panel *b*, a single school has been adjoined to each of these districts. The axiom states that this addition should not affect which district is more segregated: $X \uplus Z$ is at least as segregated as $Y \uplus Z$ if and only if X is at least as segregated as Y .

Independence (IND) Let $X, Y \in \mathcal{C}$ be two districts with equal populations and equal group distributions. Then for any district $Z \in \mathcal{C}$ that contains a single school, $X \succsim Y$ if and only if $X \uplus Z \succsim Y \uplus Z$.

Intuitively, the district $X \uplus Z$ can be thought of as comprising two “clusters”: X and Z . Define segregation *within* a cluster to be the segregation ranking of the cluster viewed in isolation. Likewise, let *between-cluster* segregation be the segregation of the district when each cluster is regarded as a single school. Independence follows from the notion that a district’s segregation should be decomposable into a weighted sum of its between-cluster and its within-cluster segregations. Why? First, between-cluster segregation is the same

¹⁰For instance, X and Y may each have 100 blacks, 1000 whites, and 50 Asians.

in each combined district in panel b . This is because X and Y have the same numbers of members of each ethnic group, so they are identical when each is considered as a single school. Moreover, segregation *within* cluster Z is clearly the same in the two combined districts. Therefore, which of the combined districts in panel b is more segregated reduces to whether cluster X is more segregated than cluster Y . This is the axiom of Independence. In Section 5.2 we show that Independence is a precondition for an index to be additively decomposable in a sense discussed by Hutchens [12].

By applying Independence repeatedly, the following lemma shows that the district Z in the axiom can actually contain any number of schools. This property will be used interchangeably with Independence.

Lemma 1 *Suppose the segregation ordering \succsim satisfies IND. Let $X, Y \in \mathcal{C}$ be two districts with equal populations and equal group distributions. Then for all districts $Z \in \mathcal{C}$ containing any number of schools, $X \succsim Y$ if and only if $X \uplus Z \succsim Y \uplus Z$.*

Proof. Let the schools of Z be enumerated: n_1, \dots, n_N . By IND, $X \succsim Y$ if and only if $X \uplus \langle n_1 \rangle \succsim Y \uplus \langle n_1 \rangle$, where $\langle n_1 \rangle$ denotes a district that consists of school n_1 alone. The districts $X' = X \uplus \langle n_1 \rangle$ and $Y' = Y \uplus \langle n_1 \rangle$ have the same size and group distribution since X and Y do. Hence, by IND, $X' \succsim Y'$ if and only if $X' \uplus \langle n_2 \rangle \succsim Y' \uplus \langle n_2 \rangle$. The result follows by repeating the same argument for schools n_3, \dots, n_N . Q.E.D.

The next axiom is the Weak School Division Property. This axiom states that a one-school district cannot become *less* segregated if the school is split into two new schools. In addition, if the new schools have identical ethnic distributions, then segregation is unchanged. Intuitively, since a single school is not segregated at all, splitting the school cannot lead to lower segregation.¹¹ And if the new schools have the same ethnic distribution, then the new district is not segregated at all, like the original district.

¹¹Our motivating example uses schools as the basic locational unit, so it ignores ability tracking and other forms of within-school segregation. Our approach could easily be used to study these phenomena by redefining basic locational unit to be the classroom or the ability group.

Weak School Division Property (WSDP) Let $X \in \mathcal{C}$ be a district consisting of a single school. Let X' be the district that results from subdividing this school into two schools, n_1 and n_2 . Then, $X' \succcurlyeq X$. Further, if n_1 and n_2 have the same group distributions (i.e., $p_g^{n_1} = p_g^{n_2}$ for all $g \in G$), then $X' \sim X$.

By combining Independence and WSDP, we can prove that the same conclusions hold if the original district X contains any number of other schools in addition to the school that is split. This is the School Division Property:

School Division Property (SDP) Let $X \in \mathcal{C}$ be any district and let n be a school in X . Let X' be the district that results from X if school n is subdivided into two schools, n_1 and n_2 . Then, $X' \succcurlyeq X$. Further, if n_1 and n_2 have the same group distributions (i.e., $p_g^{n_1} = p_g^{n_2}$ for all $g \in G$), then $X' \sim X$.

Lemma 2 *Suppose the segregation ordering \succcurlyeq satisfies Independence and the Weak School Division Property. Then \succcurlyeq also satisfies the School Division Property.*

Proof. Let Y denote the district X less the school n :

$$\begin{aligned} X &= Y \uplus \langle n \rangle \\ X' &= Y \uplus \langle n_1, n_2 \rangle. \end{aligned}$$

By WSDP, $\langle n_1, n_2 \rangle \succcurlyeq \langle n \rangle$. By IND, $Y \uplus \langle n_1, n_2 \rangle \succcurlyeq Y \uplus \langle n \rangle$. If n_1 and n_2 have the same population distribution then the symbol \succcurlyeq can be replaced by \sim . Q.E.D.

The School Division Property is related to two properties that are discussed by James and Taeuber [15] and subsequent authors. The first is *organizational equivalence*: if a school is divided into two schools that have the same group distribution, the district's level of segregation does not change. The second is the *transfer principle*. When there are two demographic groups, the transfer principle states that if a black (white) student moves from one school to another school in which the proportion of blacks (whites) is higher, then segregation in the district rises. In the case of two ethnic groups, SDP follows from

organizational equivalence and the transfer principle.¹² But while SDP applies directly with any number of groups, it is unclear what form the transfer principle should take with more than two groups.¹³

Our next axiom, Continuity, will be needed only when Group Symmetry is dropped.

Continuity (C) For any districts $X, Y, Z \in \mathcal{C}$, the sets

$$\{c \in [0, 1] : cX \uplus (1 - c)Y \succcurlyeq Z\} \text{ and } \{c \in [0, 1] : Z \succcurlyeq cX \uplus (1 - c)Y\}$$

are closed.

Our final axiom states that there exist two districts, one strictly more segregated than the other. It is needed to rule out the trivial segregation ordering.

Nontriviality (N) There exist districts $X, Y \in \mathcal{C}$ such that $X \succ Y$.

5 Results

We first show that the asymmetric Atkinson indices satisfy all of the axioms except Group Symmetry, which is satisfied by the symmetric Atkinson index.

Proposition 1 *Let $\mathbf{w} = (w_1, \dots, w_K)$ be a list of K non-negative weights that add up to one. The segregation ordering represented by the asymmetric Atkinson index $A_{\mathbf{w}}$ satisfies SI, IND, WSDP, N, and C. The segregation order represented by the Atkinson index A also satisfies GS.*

¹²Proof available on request.

¹³For instance, suppose a black student moves to a school that has higher proportions of both blacks and Asians but fewer whites. Since there are more blacks, one might argue (using the transfer principle) that segregation has gone up. On the other hand, blacks are now more integrated with Asians. One attempt to overcome this difficulty appears in Reardon and Firebaugh [21].

Proof. That the asymmetric Atkinson order satisfies N is obvious. It also satisfies C, since it is represented by the continuous function $A_{\mathbf{w}}$. The fact that the Atkinson ordering satisfies SI follows from the fact that for any positive scalar α , $t_g^n = \frac{T_g^n}{T_g} = \frac{\alpha T_g^n}{\alpha T_g}$. We now show that it satisfies IND and WSDP.

IND Let $X, Y \in \mathcal{C}$ be two districts with the same group distributions, and the same total populations, and let $Z \in \mathcal{C}$ be another district. We wish to show that $A_{\mathbf{w}}(X) \geq A_{\mathbf{w}}(Y)$ if and only if $A_{\mathbf{w}}(X \uplus Z) \geq A_{\mathbf{w}}(Y \uplus Z)$. Let $\gamma_g = \frac{T_g(X)}{T_g(X \uplus Z)} = \frac{T_g(Y)}{T_g(Y \uplus Z)}$ and $\eta_g = \frac{T_g(Z)}{T_g(X \uplus Z)} = \frac{T_g(Z)}{T_g(Y \uplus Z)}$. Note that a proportion $t_g^n \gamma_g$ of group- g students of the district $X \uplus Z$ attend school $n \in N(X)$. Likewise, a proportion $t_g^n \eta_g$ of group- g students of the district $X \uplus Z$ attend school $n \in N(Z)$. Analogous statements are true for $Y \uplus Z$. Accordingly,

$$\begin{aligned}
A_{\mathbf{w}}(X \uplus Z) &\geq A_{\mathbf{w}}(Y \uplus Z) \\
&\Leftrightarrow \sum_{n \in N(X)} \left(\prod_{g \in G} (t_g^n \gamma_g)^{w_g} \right) + \sum_{n \in N(Z)} \left(\prod_{g \in G} (t_g^n \eta_g)^{w_g} \right) \\
&\leq \sum_{n \in N(Y)} \left(\prod_{g \in G} (t_g^n \gamma_g)^{w_g} \right) + \sum_{n \in N(Z)} \left(\prod_{g \in G} (t_g^n \eta_g)^{w_g} \right) \\
&\Leftrightarrow \sum_{n \in N(X)} \left(\prod_{g \in G} (t_g^n \gamma_g)^{w_g} \right) \leq \sum_{n \in N(Y)} \left(\prod_{g \in G} (t_g^n \gamma_g)^{w_g} \right) \\
&\Leftrightarrow \left(\prod_{g \in G} (\gamma_g)^{w_g} \right) \sum_{n \in N(X)} \prod_{g \in G} (t_g^n)^{w_g} \leq \left(\prod_{g \in G} (\gamma_g)^{w_g} \right) \sum_{n \in N(Y)} \prod_{g \in G} (t_g^n)^{w_g} \\
&\Leftrightarrow \sum_{n \in N(X)} \prod_{g \in G} (t_g^n)^{w_g} \leq \sum_{n \in N(Y)} \prod_{g \in G} (t_g^n)^{w_g} \\
&\Leftrightarrow A_{\mathbf{w}}(X) \geq A_{\mathbf{w}}(Y)
\end{aligned}$$

WSDP Let X be a district with a single school and let $X' = \langle (t_g)_{g \in G}, (1 - t_g)_{g \in G} \rangle$ the district that results from dividing X into two schools. Then, since $A_{\mathbf{w}}$ maps districts to the unit interval, $A_{\mathbf{w}}(X') \geq 0 = A_{\mathbf{w}}(X)$. Further, if the two schools of X' have the same group distribution, then $t_g = t_{g'}$ for all $g, g' \in G$, then $A_{\mathbf{w}}(X') = 1 - \prod_{g \in G} t_g^{w_g} - \prod_{g \in G} (1 - t_g)^{w_g} = 0$ since the weights w_g add up to one.

GS Since the geometric average is a symmetric function, A satisfies GS.

Q.E.D.

The next two theorems are the main results of our paper. Theorem 1 states that our set of axioms, less Group Symmetry, fully characterizes the family of asymmetric Atkinson orderings.

Theorem 1 *Let \succsim be a segregation ordering on \mathcal{C} that satisfies SI, WSDP, IND, N, and C. There are fixed weights $w_g \geq 0$ for $g = 1, \dots, K$, adding up to one, such that \succsim is represented by the asymmetric Atkinson index $A_{\mathbf{w}}(X)$.*

An easy implication of Theorem 1 is that if the requirement of Group Symmetry is added, then the weights w_g must all be equal. Hence, the symmetric Atkinson ordering is the unique ordering that satisfies this larger set of axioms. It turns out that this is still true if Continuity is then removed from the set. This is the following result.

Theorem 2 *The Atkinson ordering on \mathcal{C} is the only ordering that satisfies GS, SI, WSDP, IND, and N.*

5.1 Independence of the Axioms

Are the axioms in Theorems 1 and 2 independent of each other? In this section, we show that they are: for each of the axioms in each of the two theorems, there is an index that violates it yet that satisfies the other axioms. Consequently, all of the axioms are needed for our results to hold.

We first define a new segregation ordering. For any two different vectors \mathbf{w} and \mathbf{w}' of group weights (each summing to one), consider the following lexicographic ordering:

$$X \succ_{\mathbf{w}, \mathbf{w}'} Y \text{ iff } \begin{cases} A_{\mathbf{w}}(X) > A_{\mathbf{w}}(Y) \\ \text{or} \\ A_{\mathbf{w}}(X) = A_{\mathbf{w}}(Y) \text{ and } A_{\mathbf{w}'}(X) \geq A_{\mathbf{w}'}(Y) \end{cases}$$

This ordering first uses the Atkinson index with weights \mathbf{w} to rank districts. Any “ties” are broken using the Atkinson index with weights \mathbf{w}' . The following proposition uses this index and the other indices defined in Section 3 to show that our axioms are independent of each other.

Proposition 2 *The axioms SI, WSDP, IND, N, and C are independent of each other, as are the axioms GS, SI, WSDP, IND, and N. In particular:*

- *The symmetric Atkinson index $A(X)$ satisfies all the axioms;*
- *any asymmetric Atkinson index with unequal weights satisfies all axioms but Group Symmetry;*
- *the Mutual Information index satisfies all axioms but Scale Invariance;*
- *$1 - A(X)$ satisfies all axioms but the Weak School Division Property;*
- *the Unweighted Dissimilarity index satisfies all axioms but Independence;*
- *the trivial index, which ranks all districts as equally segregated, satisfies all axioms but Nontriviality;*
- *the lexicographic index $\succ_{\mathbf{w}, \mathbf{w}'}$, for weights $\mathbf{w} \neq \mathbf{w}'$, satisfies all axioms but Group Symmetry and Continuity (so C is independent of SI, WSDP, IND, N).*

This proposition is summarized in Table 1. A check mark indicates that an index satisfies a given axiom; an \times indicates that it does not.

5.2 Additive Decomposability

It is often necessary to study segregation at several levels simultaneously. For instance, one may be interested in how much of the classroom segregation in a district is due to residential segregation, which tends to cause segregation between schools, and how much is due to ability tracking, which tends to create segregation between the classrooms of a given school.

	GS	SI	WSDP	IND	N	C
Symmetric Atkinson: $A(X)$	✓	✓	✓	✓	✓	✓
$A_{\mathbf{w}}(X)$ for $\mathbf{w} \neq (1/K, \dots, 1/K)$	×	✓	✓	✓	✓	✓
Mutual Information $M(X)$	✓	×	✓	✓	✓	✓
$1 - A(X)$	✓	✓	×	✓	✓	✓
Unweighted Dissimilarity: $D^U(X)$	✓	✓	✓	×	✓	✓
Trivial index	✓	✓	✓	✓	×	✓
Lexicographic $\succ_{\mathbf{w}, \mathbf{w}'}$ for $\mathbf{w} \neq \mathbf{w}'$	×	✓	✓	✓	✓	×

Table 1: Independence of the axioms.

As a first step in studying this issue, one might wish to write districtwide segregation as the sum of between-school segregation and within-school (between-classroom) segregation. In this section we show that only indices whose underlying orderings satisfy the axiom of Independence can be decomposed in this way. This includes the Atkinson indices but not the Unweighted Dissimilarity index.

The following notion of additive decomposability is due to Hutchens [12].¹⁴ For any district Z , let the lower-case letter z denote the one-school district that results from combining the students of Z into a single school. We say that the segregation index S is *additively decomposable* if, for any (nonempty) districts X and Y ,

$$S(X \uplus Y) = S(x \uplus y) + \alpha(x, y)S(X) + \beta(x, y)S(Y) \quad (5)$$

where $\alpha(x, y)$ and $\beta(x, y)$ are strictly positive. That is, the segregation of the combined district $X \uplus Y$ can be written as the sum of segregation between the districts, $S(x \uplus y)$, and the weighted sum of segregation within the districts X and Y , where the weights $\alpha(x, y)$ and $\beta(x, y)$ depend only on the total numbers of each ethnic group in X and Y and *not* on their allocation across schools within X or Y .

Proposition 3 *Suppose S is an additively decomposable segregation index. Then the ordering represented by S satisfies Independence.*

¹⁴Additive decomposability is one of the cardinal axioms in Hutchens's axiomatization of the 2-group symmetric Atkinson index (Hutchens [12]).

Proof. Let $X, Y \in \mathcal{C}$ be two districts with the same group distributions, and the same total populations. Let $Z \in \mathcal{C}$ be another district. We wish to show that $S(X) \geq S(Y)$ if and only if $S(X \uplus Z) \geq S(Y \uplus Z)$. Note that

$$\begin{aligned} S(X \uplus Z) &= S(x \uplus z) + \alpha(x, z)S(X) + \beta(x, z)S(Z) \text{ by (5)} \\ &= S(y \uplus z) + \alpha(y, z)S(X) + \beta(y, z)S(Z) \text{ since } x = y \end{aligned}$$

while $S(Y \uplus Z) = S(y \uplus z) + \alpha(y, z)S(Y) + \beta(y, z)S(Z)$ by (5). Since $\alpha(x, y) > 0$ by assumption, $S(X \uplus Z) - S(Y \uplus Z)$ is proportional to $S(X) - S(Y)$. Q.E.D.

The Atkinson index with weights $\mathbf{w} = (w_1, \dots, w_K)$ satisfies (5). More generally, let $Z = X_1 \uplus \dots \uplus X_N$, where each X_i is a district. Then it is straightforward to verify that

$$A_{\mathbf{w}}(Z) = A_{\mathbf{w}}(x_1 \uplus \dots \uplus x_N) + \sum_{i=1}^N \alpha_i A_{\mathbf{w}}(X_i)$$

where x_i is the district that results from combining the students in X_i into a single school and $\alpha_i = \prod_{g \in G} \left(\frac{T_g(x_i)}{T_g(z)} \right)^{w_g}$.

Frankel and Volij [10] discuss a stronger type of additive separability, in which the weight α_i equals the proportion of students who are in district i . While intuitive, this stronger property is not satisfied by the Atkinson indices or, indeed, by any of the other common segregation indices (Frankel and Volij [10]).

6 Conclusion

In this paper we have provided an axiomatic justification for the Atkinson family of segregation orderings using a parsimonious set of axioms. We have shown that the ordering represented by the symmetric Atkinson index is the only (nontrivial) segregation ordering that satisfies Symmetry, Scale Invariance, the Weak School Division Property, and Independence. We also showed that a (nontrivial) segregation ordering is represented by an asymmetric Atkinson index if and only if it satisfies Scale Invariance, the Weak School Division Property, Independence, and a technical continuity property.

Our results provide a rigorous justification for using these segregation indices in applied contexts where our axioms are suitable. Researchers can use the Atkinson indices to study segregation in the knowledge that their basic properties are simple and thoroughly understood. This is especially important since the most popular segregation index, the Dissimilarity Index (in either its weighted or unweighted form), has so far eluded an axiomatic characterization.

Our indices differ from the original Atkinson indices in two ways, one substantive and the other less so. First, we have generalized them to any number of ethnic groups. This allows their use in a general multiracial context. Second, for tractability, we have followed Hutchens [13] in using the index $A_{\mathbf{w}}(X)$ rather than the original Atkinson index of James and Taeuber [15], which equals $1 - (1 - A_{\mathbf{w}}(X))^2$. Since orderings are preserved under increasing transformations, our results apply to both versions.

6.1 Variable Number of Groups

We restrict attention to segregation orderings that rank districts with the same given number of (nonempty) ethnic groups. If one district has two groups and another three, we do not require the ordering to rank them. In fact, none of our axioms have any bite in this situation. If we wanted to extend our characterization result to the class of all districts, we would need to add an axiom restricting the way the ordering ranks districts with different numbers of groups. Frankel and Volij [10] address this issue and look for an alternative segregation measure that can be used to compare districts with different numbers of ethnic groups. Specifically, they replace Scale Invariance with a new axiom, the Group Division Property: if a given group is subdivided into two subgroups that have the same distribution across schools, then the segregation of the district should not change. In order to obtain a unique measure, they also strengthen the independence requirement. The resulting unique measure is the Mutual Information index.

The weight assigned by the Mutual Information index to a given group depends on the relative size of that group in its district. This means that if one ethnic group experiences

relative growth in a district, while maintaining its distribution across schools, the weight given by the Mutual Information index to this group will grow. For instance, if this group is relatively isolated from other groups, the Mutual Information index will tend to rise.

For researchers who dislike this property, the present paper provides an alternative: they can use an asymmetric Atkinson index that gives less weight to smaller groups. Like the Mutual Information index, such an index would be relatively insensitive to the distribution of the small groups across schools. However, it would also be Scale Invariant, and thus invariant to population growth rates that differ by ethnicity.

What weight should each group receive? One approach is to let a group’s weight equal its proportion in the universe of districts under consideration (e.g., the state or country). If longitudinal comparisons are being made, one could use an average of the group’s proportion over the different years. This could be weighted by the total population in each year or not, depending on whether or not one wants to give each year or each student equal importance in determining the group weights. While we do not provide an axiomatic justification of such an approach, it has an intuitive interpretation: a group’s importance is proportional to its relative size.

6.2 Related Literature

The first to study segregation axiomatically was Philipson [20]. He provides an axiomatic characterization of a large family of segregation orderings that have an additively separable representation. The representation consists of a weighted average of a function that depends on a school’s demographic distribution only.

The papers that are most closely related to ours are Hutchens [12, 13]. These papers study the measurement of segregation in the case of two demographic groups. Hutchens [12] characterizes the family of indices that satisfy a set of basic properties. Hutchens [13] strengthens one axiom and obtains a unique segregation index, which equals our symmetric Atkinson index in the case of two demographic groups. While we assume properties of the underlying segregation ordering, Hutchens follows the inequality literature (e.g., Shorrocks

[22, 23]) by imposing restrictions directly on the segregation index. Chakravarty and Silber [3] use a slightly different set of axioms to characterize a set of segregation indices that includes the symmetric Atkinson index.

Another related paper is Echenique and Fryer [9]. They use data on individuals' social networks to measure the strength of an individual's isolation from members of other demographic groups. Echenique and Fryer's characterize their segregation index using cardinal axioms.

A Proofs

For any district X and any nonnegative constant c , let cX denote the district that results from multiplying the number of members of each group in each school of X by c . For any district X and any vector of nonnegative scalars $\vec{\alpha} = (\alpha_g)_{g \in G}$, let $\vec{\alpha} * X$ denote the district in which the number of members of group g in school n is $\alpha_g T_g^n$. For example, if $X = \langle (1, 2), (3, 4) \rangle$, and $\vec{\alpha} = (2, 3)$, then $\vec{\alpha} * X = \langle (2, 6), (6, 12) \rangle$. We sometimes apply the same operation to individual schools; e.g., $\vec{\alpha} * (1, 2) = (2, 6)$.

We first state and prove some preliminary lemmas.

Lemma 3 *Let \succsim be a segregation ordering on \mathcal{C} that satisfies SDP and SI.*

1. *All districts in which every school is representative have the same degree of segregation under \succsim .*
2. *Any district in which every school is representative is weakly less segregated under \succsim than any district in which some school is unrepresentative.*

Proof.

1. Consider any district Y in which every school is representative. Number the schools $1, \dots, N$. For each $i = 1, \dots, N$, let Y_i be the district that results from Y when the first i schools of Y are combined into a single school. By SDP, for each $i = 1, \dots, N - 1$,

$Y_i \sim Y_{i+1}$. Hence, by transitivity, $Y = Y_1 \sim Y_N$. Y_N contains a single school. But by SI, any district with a single school is as segregated as any other district with a single school.

2. Let Y be a district in which every school is representative and consider any district X in which at least one school is unrepresentative. The above reasoning yields $X \succcurlyeq X_N$. X_N contains a single school, so it is representative. Therefore, by 1, $X \succcurlyeq Y$.

Q.E.D.

Lemma 4 *Let \succcurlyeq be a segregation ordering on \mathcal{C} that satisfies SDP and SI. All completely segregated districts have the same degree of segregation under \succcurlyeq , and are weakly more segregated than any district in which any school is mixed.*

Proof. Consider a completely segregated district X . Let X' be the district that results from X when, for each group $g \in G$, all schools that contain only members of group g are combined into a single school. (X' thus consists of K schools, each of which contains all the members of a single group.) By iteratively applying SDP, $X \sim X'$. By SI, X' is as segregated as any other district that consists of K schools, each of which contains all the members of a single group. This implies that all completely segregated districts have the same degree of segregation.

Now any district that has at least one mixed school can be converted into a completely segregated district by dividing each school n into K distinct schools, each of which includes all and only the members of a single group. By SDP, this procedure results in a weakly more segregated district. Q.E.D.

Let \underline{X} be a district with K groups of unit size who all attend in the same school: $\underline{X} = \langle \underbrace{(1, 1, \dots, 1)}_{K \text{ groups}} \rangle$. Let \overline{X} be a district with K groups of unit size who all attend separate schools:

$$\overline{X} = \left\langle \underbrace{\left(\underbrace{1, 0, \dots, 0}_{K \text{ groups}}, \underbrace{0, 1, 0, \dots, 0}_{K \text{ groups}}, \dots, \underbrace{0, \dots, 0, 1}_{K \text{ groups}} \right)}_{K \text{ schools}} \right\rangle.$$

We say that a school is a *ghetto* if all its students belong to the same group. We first state and prove some auxiliary results about districts with a single non-ghetto school. For any scalar α , let $X(\alpha)$ denote the district $\alpha \underline{X} \uplus (1 - \alpha) \overline{X}$. City $X(\alpha)$ contains one school with α students of each group, and K ghettos, each with $1 - \alpha$ students. Similarly, for any vector $\mathbf{t} = (t_1, \dots, t_K) \in [0, 1]^K$, let $X(\mathbf{t})$ denote the district

$$\mathbf{t} * \underline{X} \uplus (1 - \mathbf{t}) * \overline{X} = \langle \mathbf{t}, (1 - t_1, 0, \dots, 0), (0, \dots, 0, 1 - t_K) \rangle$$

City $X(\mathbf{t})$ consists of the non-ghetto school \mathbf{t} , and for each group g , one ghetto with $1 - t_g$ students of group g .

Lemma 5 *Let \succsim be a segregation ordering on \mathcal{C} that satisfies SDP, IND, N, and SI. Then*

1. $\overline{X} \succ \underline{X}$;
2. for any $\alpha, \beta \in [0, 1]$, $\alpha > \beta$, $X(\beta) \succ X(\alpha)$.

Proof.

1. By N, there exist districts X and Y such that $X \succ Y$. By lemmas 3 and 4, $\overline{X} \succsim X \succ Y \succsim \underline{X}$, so $\overline{X} \succ \underline{X}$.
2. By part 1 and SI, $(\alpha - \beta) \overline{X} \succ (\alpha - \beta) \underline{X}$. Since the numbers of members of each group are equal in district \overline{X} and in \underline{X} , they are also equal in district $(\alpha - \beta) \overline{X}$ and in $(\alpha - \beta) \underline{X}$. So by IND,

$$\beta \underline{X} \uplus (\alpha - \beta) \overline{X} \uplus (1 - \alpha) \overline{X} \succ \beta \underline{X} \uplus (\alpha - \beta) \underline{X} \uplus (1 - \alpha) \overline{X}.$$

The result follows from the fact that, by SDP,

$$\beta \underline{X} \uplus (\alpha - \beta) \overline{X} \uplus (1 - \alpha) \overline{X} \sim \beta \underline{X} \uplus (1 - \beta) \overline{X}$$

and

$$\beta \overline{X} \uplus (\alpha - \beta) \underline{X} \uplus (1 - \alpha) \underline{X} \sim \alpha \underline{X} \uplus (1 - \alpha) \overline{X}.$$

Q.E.D.

Claim 1 Lemma 6 *Let $\mathbf{t}, \mathbf{v} \in [0, 1]^K$, such that $\mathbf{t} \leq \mathbf{v}$. Then, $X(\mathbf{t}) \succcurlyeq X(\mathbf{v})$. If $\mathbf{t} = (t_1, \dots, t_K) \in (0, 1)^K$ then $\overline{X} \succ X(\mathbf{t}) \succ \underline{X}$.*

Proof. Let $\mathbf{t}, \mathbf{v} \in [0, 1]^K$, such that $\mathbf{t} \leq \mathbf{v}$. Applying SDP twice, we obtain

$$\begin{aligned} X(\mathbf{t}) &= \mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X} \\ &\sim \mathbf{t} * \underline{X} \uplus (\mathbf{v} - \mathbf{t}) * \overline{X} \uplus (\mathbf{1} - \mathbf{v}) * \overline{X} \\ &\succcurlyeq \mathbf{v} * \underline{X} \uplus (\mathbf{1} - \mathbf{v}) * \overline{X} = X(\mathbf{v}). \end{aligned}$$

Assume now that $\mathbf{t} = (t_1, \dots, t_K) \in (0, 1)^K$, and let $\bar{t} = \max\{t_1, \dots, t_K\}$, $\underline{t} = \min\{t_1, \dots, t_K\}$. Then,

$$\begin{aligned} \overline{X} &\succ X(\underline{t}) \text{ by Lemma 5} \\ &\succcurlyeq X(\mathbf{t}) \text{ since } (\underline{t}, \dots, \underline{t}) \leq \mathbf{t} \\ &\succcurlyeq X(\bar{t}) \text{ since } \mathbf{t} \leq (\bar{t}, \dots, \bar{t}) \\ &\succ \underline{X} \text{ by Lemma 5.} \end{aligned}$$

Q.E.D.

Lemma 7 *For any two vectors $\mathbf{t}, \mathbf{v} \in [0, 1]^K$ and for any $\gamma \in (0, 1]$,*

1. $\mathbf{v} * X(\mathbf{t}) \uplus (\mathbf{1} - \mathbf{v}) * \overline{X} \sim X(\mathbf{v} * \mathbf{t})$
2. $\gamma X(\mathbf{t}) \uplus (1 - \gamma) \overline{X} \sim X(\gamma \mathbf{t})$
3. *If for some $\alpha \in [0, 1]$, $X(\mathbf{t}) \sim X(\alpha)$, then $X(\mathbf{v} * \mathbf{t}) \sim X(\alpha \mathbf{v})$*
4. *If for some $\alpha \in [0, 1]$, $X(\mathbf{t}) \sim X(\alpha)$, then $X(\gamma \mathbf{t}) \sim X(\gamma \alpha)$*

Proof.

1. By definition of $X(\mathbf{t})$ and by SDP,

$$\begin{aligned} \mathbf{v} * X(\mathbf{t}) \uplus (\mathbf{1} - \mathbf{v}) * \overline{X} &= \mathbf{v} * (\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X}) \uplus (\mathbf{1} - \mathbf{v}) * \overline{X} \\ &\sim (\mathbf{v} * \mathbf{t}) * \underline{X} \uplus (\mathbf{1} - \mathbf{v} * \mathbf{t}) * \overline{X} \\ &= X(\mathbf{v} * \mathbf{t}). \end{aligned}$$

2. The proof is analogous to the previous one.

3. Now, if for some $\alpha \in [0, 1]$, $X(\mathbf{t}) \sim X(\alpha)$, then, by SI and IND

$$\mathbf{v} * X(\mathbf{t}) \uplus (\mathbf{1} - \mathbf{v}) * \overline{X} \sim \mathbf{v} * X(\alpha) \uplus (\mathbf{1} - \mathbf{v}) * \overline{X}$$

which, by the previous steps, implies $X(\mathbf{v} * \mathbf{t}) \sim X(\alpha \mathbf{v})$.

4. The proof is analogous to the previous one.

Q.E.D.

A.1 Proof of Theorem 1

Let \succsim be a segregation ordering on \mathcal{C} that satisfies C, SDP, IND, N, and SI. We first build an index that represents \succsim . Later we show that the index has the requisite form.

Lemma 8 *For any district X , there is a unique $\alpha_X \in [0, 1]$ such that $X \sim X(\alpha_X)$.*

Proof. By C, $\{\alpha \in [0, 1] : \alpha \underline{X} \uplus (1 - \alpha) \overline{X} \succsim X\}$ and $\{\alpha \in [0, 1] : X \succsim \alpha \underline{X} \uplus (1 - \alpha) \overline{X}\}$ are closed sets. Any α_X satisfies $X \sim X(\alpha_X)$ if and only if it is in the intersection of these two sets. The sets are each nonempty by Lemmas 3 and 4. Their union is the whole unit interval since \succsim is complete. Since the interval $[0, 1]$ is connected, the intersection of the two sets must be nonempty. By Lemma 5, their intersection cannot contain more than one element. Thus, their intersection contains a single element α_X . Q.E.D.

Let X and Y be two districts, and let α_X , and α_Y be the respective scalars identified in Lemma 8. Then, by Lemma 5, $X \succcurlyeq Y$ if and only if $1 - \alpha_X > 1 - \alpha_Y$, which implies that the index $S : \mathcal{C} \rightarrow [0, 1]$ defined by $S(Z) = 1 - \alpha_Z$ represents \succcurlyeq .

We will now show that the index S has the requisite form.

Proposition 4 *For each group g there is a fixed constant $w_g \geq 0$ such that for any $\beta \in (0, 1]$,*

$$X((1, \dots, 1, \beta, 1, \dots, 1)) \sim X(\beta^{w_g})$$

where $(1, \dots, 1, \beta, 1, \dots, 1)$ is a vector with β in the g th place and ones elsewhere.

Proof. Let \mathfrak{R}_+ denote the nonnegative reals. For any scalar β and any group g , let β_g denote the vector $(1, \dots, 1, \beta, 1, \dots, 1)$ with β in the g th place and ones elsewhere. Similarly, for all $q \in \mathfrak{R}_+$, β_g^q denotes the vector $(1, \dots, 1, \beta^q, 1, \dots, 1)$. We prove the proposition by means of the following two lemmas.

Lemma 9 *Fix some $\beta \in (0, 1]$, and let δ be the unique scalar such that $X((1, \dots, 1, \beta, 1, \dots, 1)) \sim X(\delta)$. Then, for all $q \in \mathfrak{R}_+$*

$$X(\beta_g^q) \sim X(\delta^q) \tag{6}$$

Proof. We first show that δ satisfies (6) for all natural numbers q . By assumption δ satisfies (6) for $q = 1$. Assume that for some natural n

$$X(\beta_g^n) \sim X(\delta^n) \tag{7}$$

Then

$$\begin{aligned} X(\beta_g^{n+1}) &\sim X(\beta_g * (\beta_g^n)) \\ &\sim X(\delta(\beta_g^n)) && \text{by Lemma 7} \\ &\sim \delta X(\beta_g^n) \uplus (1 - \delta) \overline{X} && \text{by Lemma 7} \\ &\sim \delta X(\delta^n) \uplus (1 - \delta) \overline{X} && \text{by (7) and IND} \\ &\sim X(\delta^{n+1}) && \text{by Lemma 7} \end{aligned}$$

This shows that (7) holds for all naturals.

We now show that δ satisfies (6) for all rational numbers. Let $q = n/m$ be a rational number. Then, $(\beta^n)_g = (\beta^q)_g * (\beta^m)_g$, and hence,

$$X(\beta_g^n) \sim X(\beta_g^q * \beta_g^m).$$

By Lemma 8, there is a κ such that

$$X(\beta_g^q) \sim X(\kappa) \tag{8}$$

Therefore, by Lemma 7 and since (7) holds for all naturals,

$$\begin{aligned} X(\beta_g^n) &\sim X(\kappa \beta_g^m) \\ &\sim X(\kappa \delta^m) \end{aligned}$$

On the other hand, since equation (7) holds for all naturals

$$X(\beta_g^n) \sim X(\delta^n)$$

hence $X(\kappa \delta^m) \sim X(\delta^n)$ which, by Lemma 5, implies $\kappa = \delta^q$. Substituting into equation (8), we get

$$X(\beta_g^q) \sim X(\delta^q).$$

It remains to show that the statement of the lemma holds for all reals. Let $r \in \mathfrak{R}_+$. By C, the sets

$$S_1 = \left\{ c \in [0, 1] : X((\beta^r)_g) \succcurlyeq X(c) \right\} \text{ and } S_2 = \left\{ c \in [0, 1] : X((\beta^r)_g) \preccurlyeq X(c) \right\}$$

are both closed. Let $\{q_n\}$ be a sequence rational numbers such that $q_n \geq r$ for all n that converge to r . Since $\beta \leq 1$, $\beta^{q_n} \leq \beta^r$, so, by Lemma 6, $X(\beta_g^{q_n}) \succcurlyeq X(\beta_g^r)$ for all n . Since for each n , q_n is rational, $X(\beta_g^{q_n}) \sim X(\delta^{q_n})$. As a result we obtain $X(\delta^{q_n}) \succcurlyeq X(\beta_g^r)$. So $\delta^{q_n} \in S_2$ for all n . Since S_2 is closed and δ^{q_n} converges to δ^r , we conclude that $\delta^r \in S_2$. A similar argument shows that $\delta^r \in S_1$ as well. As a result $X(\beta_g^r) \sim X(\delta^r)$. Q.E.D.

Lemma 10 *For all g , there is a fixed constant $w_g \geq 0$ such that for any $\beta \in (0, 1]$, the δ given by Lemma 9 is $\delta = \beta^{w_g}$.*

Proof. Pick some $\beta \in (0, 1)$ and let δ the scalar identified in Lemma 9. We must have $\beta \leq \delta$. Otherwise we would have $X(\beta_g) \sim X(\delta) \succ X(\beta)$ contradicting Lemma 6. Let w_g satisfy $\delta = \beta^{w_g}$. Since $\beta \leq \delta < 1$, $w_g \geq 0$. Now consider any $\beta' \in (0, 1]$. Let r be such that $\beta' = \beta^r$. By Lemma 9,

$$X(\beta'^q_g) = X(\beta^{rq}_g) \sim X(\delta^{rq}) = X([\delta^r]^q) \text{ for all } q \in \mathbb{R}_+$$

which shows that the δ' corresponding to β' is just δ^r . Hence, $\delta' = \delta^r = (\beta^{w_g})^r = (\beta^r)^{w_g} = \beta'^{w_g}$. Q.E.D.

This ends the proof of Proposition 4 Q.E.D.

Proposition 5 *There are fixed, non-negative weights $w_g \geq 0$ for $g = 1, \dots, K$ such that for any $\mathbf{t} \in [0, 1]^K$ the unique $\alpha \in [0, 1]$ that satisfies $X(\mathbf{t}) \sim X(\alpha)$ is given by $\prod_{g=1}^K (t_g)^{w_g}$. Further, the weights add up to one.*

Proof. Case 1: $\mathbf{t} \in (0, 1]^K$.

Let $\mathbf{t} = (t_1, \dots, t_K) \in (0, 1]^K$. By Proposition 4,

$$X((1, 1, \dots, 1, t_g, 1, \dots, 1)) \sim X(t_g^{w_g}) \text{ for all } g = 1, \dots, K.$$

Note that $\mathbf{t} = (t_1, 1, \dots, 1) * (1, t_2, 1, \dots, 1) * \dots * (1, \dots, 1, t_k)$. Then, repeated applications of Lemma 7 then yields

$$X(\mathbf{t}) = X\left(\prod_{g=1}^K t_g^{w_g}\right).$$

In order to complete the proof of case 1, we need to show that the weights w_g add up to one. Consider the district $X = X(\alpha)$ where $\alpha \in (0, 1)$. By the previous conclusion $X \sim X\left(\prod_{g=1}^K \alpha^{w_g}\right)$. By Lemma 5 $\left(\prod_{g=1}^K \alpha^{w_g}\right) = \alpha$ which implies that the weights w_g add up to one.

Case 2: $\mathbf{t} \in [0, 1]^K \setminus (0, 1]^K$.

By Lemma 8 there is an $\alpha \in [0, 1]$ such that $X(\mathbf{t}) \sim X(\alpha)$. We need to show that $\alpha > 0$. Let $\mathbf{t}(\varepsilon) = (t_1(\varepsilon), \dots, t_K(\varepsilon))$ be the school that results from \mathbf{t} after replacing the 0 components by $\varepsilon > 0$. Since $\mathbf{t} \in (0, 1]^K$, by Case 1, $X(\mathbf{t}(\varepsilon)) \sim X(\alpha(\varepsilon))$ where $\alpha(\varepsilon) = \prod_{g=1}^K t_g(\varepsilon)^{w_g}$. By Lemma 6, $X(\mathbf{t}) \succsim X(\mathbf{t}(\varepsilon))$ which implies, $X(\alpha) \succsim X(\alpha(\varepsilon))$. By Lemma 5, $\alpha(\varepsilon) > \alpha \geq 0$. Since $\alpha(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$, we obtain that $\alpha = 0$. Q.E.D.

We now show that the statement of the theorem holds for districts with two non-ghetto schools.

Proposition 6 *Let $\mathbf{t}^1, \mathbf{t}^2 \in [0, 1]^K$ and let $X = \langle \mathbf{t}^1, \mathbf{t}^2, (1 - t_1^1 - t_1^2, 0, \dots, 0), \dots, (0, \dots, 0, 1 - t_K^1 - t_K^2) \rangle$ be a district. There is a unique $\alpha_X \in [0, 1]$ that satisfies $X \sim X(\alpha_X)$. It is given by $\alpha_X = \prod_{g=1}^K (t_g^1)^{w_g} + \prod_{g=1}^K (t_g^2)^{w_g}$, where the weights w_g are those found in Proposition 5.*

Proof. Uniqueness of α_X follow from Lemma 5, so it is enough to show that $\alpha_X = \prod_{g=1}^K (t_g^1)^{w_g} + \prod_{g=1}^K (t_g^2)^{w_g}$ satisfies $X \sim X(\alpha_X)$. Assume first that $t_g^i \leq 1/2$ for $i = 1, 2$ and $g = 1, \dots, K$. First suppose that $t_g^i = 0$ for some i and g . Assume WLOG that $t_1^2 = 0$. Then by Proposition 5 and SI,

$$\langle \mathbf{t}^2, (1 - t_1^1 - t_1^2, 0, \dots, 0), \dots, (0, \dots, 0, 1 - t_K^1 - t_K^2) \rangle \sim \langle (1 - t_1^1, 0, \dots, 0), \dots, (0, \dots, 0, 1 - t_K^1) \rangle$$

so by IND, $X \sim X(\mathbf{t}^1)$. The result then follows from Proposition 5.

Now suppose that $\mathbf{t}^1, \mathbf{t}^2 \in (0, 1]^K$. Assume WLOG that $\prod_{g=1}^K (t_g^1)^{w_g} \leq \prod_{g=1}^K (t_g^2)^{w_g}$. Define $\tilde{t}_g^i = t_g^i / (1 - t_g^2)$ for $g = 1, \dots, K$ and $i = 1, 2$. Note that $\prod_{g=1}^K (\tilde{t}_g^1)^{w_g} \leq \prod_{g=1}^K (\tilde{t}_g^2)^{w_g}$. Define $\tau = \prod_{g=1}^K \left(\frac{t_g^1}{t_g^2} \right)^{w_g} = \prod_{g=1}^K \left(\frac{\tilde{t}_g^1}{\tilde{t}_g^2} \right)^{w_g} \leq 1$. We can write

$$X = \langle \mathbf{t}^1, (1 - t_1^1 - t_1^2, 0, \dots, 0), (0, 1 - t_2^1 - t_2^2, 0, \dots, 0), \dots, (0, \dots, 0, 1 - t_K^1 - t_K^2) \rangle \uplus \langle \mathbf{t}^2 \rangle.$$

By SI

$$X \sim Y \uplus \left\langle \left(\tilde{t}_1^2, \dots, \tilde{t}_K^2 \right) \right\rangle \quad (9)$$

where¹⁵

$$\begin{aligned} Y &= \left\langle \left(\widetilde{t}_1^1, \dots, \widetilde{t}_K^1 \right), \left(1 - \widetilde{t}_1^1, 0, \dots, 0 \right), \dots, \left(0, \dots, 0, 1 - \widetilde{t}_K^1 \right) \right\rangle \\ &= \widetilde{\mathbf{t}}^1 * \underline{X}^K \uplus (\mathbf{1} - \widetilde{\mathbf{t}}^1) * \overline{X}^K. \end{aligned}$$

By Proposition 5,

$$Y \sim \alpha_Y \underline{X} \uplus (1 - \alpha_Y) \overline{X}. \quad (10)$$

where $\alpha_Y = \prod_{g=1}^K \left(\widetilde{t}_g^1 \right)^{w_g}$. Define

$$Y' = \tau \widetilde{t}^2 * \underline{X} \uplus (\mathbf{1} - \tau \widetilde{t}^2) * \overline{X}. \quad (11)$$

We must verify that all entries in Y' are nonnegative. This holds if $\tau \widetilde{t}_g^2 \leq 1$ for all g . Since $t_g^2 \leq 1/2$ for all g , it follows that $\widetilde{t}_g^2 \leq 1$; since $\tau \leq 1$ as well, it follows that $\tau \widetilde{t}_g^2 \leq 1$. Since $\prod_{g=1}^K \left(\tau \widetilde{t}_g^2 \right)^{w_g} = \prod_{g=1}^K \left(\widetilde{t}_g^1 \right)^{w_g} = \alpha_Y$, by Proposition 5,

$$Y' \sim \alpha_Y \underline{X} \uplus (1 - \alpha_Y) \overline{X}. \quad (12)$$

It follows from (10) and (12) that $Y \sim Y'$. As a result,

$$\begin{aligned} X &\sim Y \uplus \left\langle \left(\widetilde{t}_1^2, \dots, \widetilde{t}_K^2 \right) \right\rangle && \text{by (9)} \\ &\sim Y' \uplus \left\langle \left(\widetilde{t}_1^2, \dots, \widetilde{t}_K^2 \right) \right\rangle && \text{by IND} \\ &\sim \tau \widetilde{t}^2 * \underline{X} \uplus (\mathbf{1} - \tau \widetilde{t}^2) * \overline{X} \uplus \left\langle \left(\widetilde{t}_1^2, \dots, \widetilde{t}_K^2 \right) \right\rangle && \text{by (11)} \\ &\sim (\tau + 1) \widetilde{\mathbf{t}}^2 * \underline{X} \uplus (\mathbf{1} - \tau \widetilde{\mathbf{t}}^2) * \overline{X} && \text{by SDP} \\ &\sim (\tau + 1) \mathbf{t}^2 * \underline{X} \uplus (\mathbf{1} - (\tau + 1) \mathbf{t}^2) * \overline{X} && \text{by SI and definition of } \widetilde{\mathbf{t}}^2. \end{aligned}$$

Therefore, using Proposition 5, $X \sim \alpha_X \underline{X} \uplus (1 - \alpha_X) \overline{X}$, where

$$\alpha_X = (\tau + 1) \prod_{g=1}^K \left(\widetilde{t}_g^2 \right)^{w_g} = \prod_{g=1}^K \left(\widetilde{t}_g^1 \right)^{w_g} + \prod_{g=1}^K \left(t_g^2 \right)^{w_g}.$$

¹⁵We must check that Y has no negative entries. Since X cannot have negative entries, it must be that $t_g^1 + t_g^2 \leq 1$ for all g . Since in addition $t_g^2 < 1$ for all g , it follows that $\frac{t_g^1}{1-t_g^2} \leq 1$ for all g . Hence, all entries in Y are nonnegative.

Consider now the case of general $\mathbf{t}^1, \mathbf{t}^2 \in [0, 1]^2$. Define $\hat{t}^i = \frac{1}{2}t^i$ for $i = 1, 2$. Let

$$\hat{X} = \langle \hat{\mathbf{t}}^1, \hat{\mathbf{t}}^2, (1 - \hat{t}_1^1 - \hat{t}_1^2, 0, \dots, 0), (0, 1 - \hat{t}_2^1 - \hat{t}_2^2, 0, \dots, 0), \dots, (0, \dots, 0, 1 - \hat{t}_K^1 - \hat{t}_K^2) \rangle.$$

Each entry in each vector is at most one half. By the preceding argument, there is a unique $\hat{\alpha}_X \in [0, 1]$ such that

$$\hat{X} \sim \hat{\alpha}_X \underline{X} \uplus (1 - \hat{\alpha}_X) \overline{X}. \quad (13)$$

and this unique $\hat{\alpha}_X$ is $\prod_{g=1}^K (\hat{t}_g^1)^{w_g} + \prod_{g=1}^K (\hat{t}_g^2)^{w_g}$. Further note that by SDP, $\hat{X} \sim \frac{1}{2}X \uplus \frac{1}{2}\overline{X}$. Therefore

$$\begin{aligned} \frac{1}{2}X \uplus \frac{1}{2}\overline{X} &\sim \hat{\alpha}_X \underline{X} \uplus (1 - \hat{\alpha}_X) \overline{X} \\ &\sim \frac{1}{2} (2\hat{\alpha}_X) \underline{X} \uplus (1 - \frac{1}{2} (2\hat{\alpha}_X)) \overline{X} \\ &\sim \frac{1}{2} (2\hat{\alpha}_X) \underline{X} \uplus \frac{1}{2} (1 - (2\hat{\alpha}_X)) \overline{X} \uplus \frac{1}{2} \overline{X} \end{aligned}$$

where the last line follows from SDP. Finally, by IND and SI

$$X \sim (2\hat{\alpha}_X) \underline{X} \uplus (1 - (2\hat{\alpha}_X)) \overline{X}$$

which means that the unique α_X that we are looking for is $\alpha_X = 2\hat{\alpha}_X = \prod_{g=1}^K (t_g^1)^{w_g} + \prod_{g=1}^K (t_g^2)^{w_g}$. Q.E.D.

Proposition 7 *For every district $X \in \mathcal{C}$ there is a unique $\alpha_X \in [0, 1]$ such that $X \sim \alpha_X \underline{X} \uplus (1 - \alpha_X) \overline{X}$. Further, this unique α_X is $\sum_{n \in N(X)} \prod_{g=1}^K (t_g^n)^{w_g}$, where the weights w_g are those found in Proposition 5.*

Proof. By SI it is enough to prove the statement for districts where all groups have a population measure of one. Also, by SDP we can restrict attention to districts where for each group there is at most one ghetto. The proof is by induction on the number of non-ghetto schools. Propositions 5 and 6 already show that the statement is true for districts

with at most two non-ghetto schools. Assume that the statement of the theorem holds for all districts with $m - 1$ non-ghetto schools, let

$$X = \langle \mathbf{t}^1, \dots, \mathbf{t}^m, (1 - \sum_{n=1}^m t_1^n, 0, \dots, 0), (0, 1 - \sum_{n=1}^m t_2^n, 0, \dots, 0), \dots, (0, \dots, 0, 1 - \sum_{n=1}^m t_K^n) \rangle$$

be a district with m non-ghetto schools. Then one can write

$$X = Y \uplus \langle \mathbf{t}^m \rangle$$

where Y denotes X with school \mathbf{t}^m removed. Y has $m - 1$ non-ghetto schools. By SI

$$Y \uplus \langle \mathbf{t}^m \rangle \sim \left[\left(\frac{1}{1 - t_1^m}, \dots, \frac{1}{1 - t_K^m} \right) * Y \right] \uplus \left\langle \left(\frac{t_1^m}{1 - t_1^m}, \dots, \frac{t_K^m}{1 - t_K^m} \right) \right\rangle.$$

By the induction hypothesis, $\left(\frac{1}{1 - t_1^m}, \dots, \frac{1}{1 - t_K^m} \right) * Y \sim \alpha_Y \underline{X} \uplus (1 - \alpha_Y) \overline{X}$ where

$$\alpha_Y = \sum_{n=1}^{m-1} \prod_{g=1}^K \left(\frac{t_g^n}{1 - t_g^m} \right)^{w_g}.$$

Using (in order) IND, SI, and Proposition 6,

$$\begin{aligned} & \left[\left(\frac{1}{1 - t_1^m}, \dots, \frac{1}{1 - t_K^m} \right) * Y \right] \uplus \left\langle \left(\frac{t_1^m}{1 - t_1^m}, \dots, \frac{t_K^m}{1 - t_K^m} \right) \right\rangle \\ & \sim \alpha_Y \underline{X} \uplus (1 - \alpha_Y) \overline{X} \uplus \left\langle \left(\frac{t_1^m}{1 - t_1^m}, \dots, \frac{t_K^m}{1 - t_K^m} \right) \right\rangle \\ & \sim (1 - t_1^m, \dots, 1 - t_K^m) * (\alpha_Y \underline{X} \uplus (1 - \alpha_Y) \overline{X}) \uplus \langle \mathbf{t}^m \rangle \\ & \sim \alpha_X \underline{X} \uplus (1 - \alpha_X) \overline{X} \end{aligned}$$

where

$$\begin{aligned} \alpha_X &= \prod_{g=1}^K (1 - t_g^m)^{w_g} \alpha_Y + \prod_{g=1}^K (t_g^m)^{w_g} \\ &= \prod_{g=1}^K (1 - t_g^m)^{w_g} \sum_{n=1}^{m-1} \prod_{g=1}^K \left(\frac{t_g^n}{1 - t_g^m} \right)^{w_g} + \prod_{g=1}^K (t_g^m)^{w_g} \\ &= \sum_{n=1}^m \prod_{g=1}^K (t_g^n)^{w_g}. \end{aligned}$$

Q.E.D.

This completes the proof of Theorem 1.

A.2 Proof of Theorem 2

Proposition 1 implies that the Atkinson index A satisfies all the axioms of the theorem. We now show that it is the only index to do so. We now show that any ordering that satisfies GS, SI, SDP, IND, and N on \mathcal{C} must be the Atkinson ordering. Let \succsim be such an ordering.

Proposition 8 *Let $\mathbf{t} = (t_1, \dots, t_K) \in [0, 1]^K$ and let $X = X(\mathbf{t})$. Then, there exists a unique $\alpha_X \in [0, 1]$ such that $X \sim X(\alpha_X)$. Further, this unique α_X is $\left(\prod_{g=1}^K t_g\right)^{1/K}$.*

Proof. For existence, there are two cases.

Case 1: Suppose $t_g = 0$ for some g . In this case we have to show that $\alpha_X = 0$ or, equivalently, that $X \sim \overline{X}$. By GS, we can assume w.l.o.g. that $t_1 = 0$. Therefore $\mathbf{t} = (0, t_2, t_3, \dots, t_K)$. Let σ_{12} be the permutation that relabels groups 1 and 2 into 2 and 1, respectively. Therefore, $\sigma_{12}\mathbf{t} = (t_2, 0, t_3, \dots, t_K)$. Let $\mathbf{1}$ denote a vector of K ones. By GS,

$$\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X} \sim \sigma_{12}\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \sigma_{12}\mathbf{t}) * \overline{X}.$$

For any $\beta \in (0, 1)$, let $\gamma = (\beta, 1, \dots, 1)$. By SI and IND,

$$\gamma * (\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X}) \uplus (1 - \gamma) * \overline{X} \sim \gamma * (\sigma_{12}\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \sigma_{12}\mathbf{t}) * \overline{X}) \uplus (1 - \gamma) * \overline{X}.$$

Hence, by SDP and GS,

$$\begin{aligned} (\gamma * \mathbf{t}) * \underline{X} \uplus (\mathbf{1} - \gamma * \mathbf{t}) * \overline{X} &\sim (\gamma * \sigma_{12}\mathbf{t}) * \underline{X} \uplus (\mathbf{1} - \gamma * \sigma_{12}\mathbf{t}) * \overline{X} \\ &\sim [\sigma_{12}(\gamma * \sigma_{12}\mathbf{t})] * \underline{X} \uplus (\mathbf{1} - [\sigma_{12}(\gamma * \sigma_{12}\mathbf{t})]) * \overline{X}. \end{aligned} \quad (14)$$

But note that since $(\gamma * \mathbf{t}) = \mathbf{t}$, and $\sigma_{12}(\gamma * \sigma_{12}\mathbf{t}) = (0, \beta t_2, t_3, \dots, t_K)$, we can write (14) as

$$\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X} \sim (0, \beta t_2, t_3, \dots, t_K) * \underline{X} \uplus (\mathbf{1} - (0, \beta t_2, t_3, \dots, t_K)) * \overline{X}.$$

We can repeat this procedure for t_3, \dots, t_K to obtain

$$\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X} \sim (0, \beta t_2, \beta t_3, \dots, \beta t_K) * \underline{X} \uplus (\mathbf{1} - (0, \beta t_2, \beta t_3, \dots, \beta t_K)) * \overline{X}$$

namely,

$$X \sim \beta \mathbf{t} * \underline{X} \uplus (\mathbf{1} - \beta \mathbf{t}) * \overline{X} \quad \text{for all } \beta \in (0, 1). \quad (15)$$

Now choose some constants $\beta, \beta' \in (0, 1)$, $\beta > \beta'$. It follows from (15) that

$$\beta \mathbf{t} * \underline{X} \uplus (\mathbf{1} - \beta \mathbf{t}) * \overline{X} \sim \beta' \mathbf{t} * \underline{X} \uplus (\mathbf{1} - \beta' \mathbf{t}) * \overline{X}.$$

Since $\beta \mathbf{t} = \underbrace{\beta' \mathbf{t}}_{\geq 0} + \underbrace{(\beta - \beta') \mathbf{t}}_{\geq 0}$, and $\mathbf{1} - \beta' \mathbf{t} = \underbrace{(\beta - \beta') \mathbf{t}}_{\geq 0} + \underbrace{(\mathbf{1} - \beta \mathbf{t})}_{\geq 0}$, by SDP

$$\beta' \mathbf{t} * \underline{X} \uplus (\beta - \beta') \mathbf{t} * \underline{X} \uplus (\mathbf{1} - \beta \mathbf{t}) * \overline{X} \sim \beta' \mathbf{t} * \underline{X} \uplus (\beta - \beta') \mathbf{t} * \overline{X} \uplus (\mathbf{1} - \beta \mathbf{t}) * \overline{X}$$

Note that $(\mathbf{1} - \beta \mathbf{t}) = \underbrace{(\beta - \beta')(\mathbf{1} - \mathbf{t})}_{\geq 0} + \underbrace{[(1 - \beta)\mathbf{1} + \beta'(\mathbf{1} - \mathbf{t})]}_{\geq 0}$, so we can subdivide $(\mathbf{1} - \beta \mathbf{t}) * \overline{X}$ in the above expression using SDP again and get

$$\begin{aligned} & \beta' \mathbf{t} * \underline{X} \uplus (\beta - \beta') \mathbf{t} * \underline{X} \uplus (\beta - \beta')(\mathbf{1} - \mathbf{t}) * \overline{X} \uplus [(1 - \beta)\mathbf{1} + \beta'(\mathbf{1} - \mathbf{t})] * \overline{X} \\ & \sim \beta' \mathbf{t} * \underline{X} \uplus (\beta - \beta') \mathbf{t} * \overline{X} \uplus (\beta - \beta')(\mathbf{1} - \mathbf{t}) * \overline{X} \uplus [(1 - \beta)\mathbf{1} + \beta'(\mathbf{1} - \mathbf{t})] * \overline{X}. \end{aligned}$$

By IND,

$$(\beta - \beta') \mathbf{t} * \underline{X} \uplus (\beta - \beta')(\mathbf{1} - \mathbf{t}) * \overline{X} \sim (\beta - \beta') \mathbf{t} * \overline{X} \uplus (\beta - \beta')(\mathbf{1} - \mathbf{t}) * \overline{X}.$$

Finally by definition of SI, $\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X} \sim \mathbf{t} * \overline{X} \uplus (\mathbf{1} - \mathbf{t}) * \overline{X} = \overline{X}$, as claimed.

Q.E.D.

Case 2. Suppose $t_g \in (0, 1]$ for all g . Let $\alpha = \left(\prod_{g=1}^K t_g \right)^{1/K}$, and let

$$Y = \alpha \underline{X} \uplus (1 - \alpha) \overline{X} = \langle (\alpha, \dots, \alpha), (1 - \alpha, 0, \dots, 0), (0, 1 - \alpha, 0, \dots, 0), \dots, (0, \dots, 0, 1 - \alpha) \rangle.$$

We shall show that $X \sim Y$ and therefore that α is the α_X we are looking for.

Let $\gamma_1 \in (0, 1)$. For $g = 2, \dots, K$, define $\gamma_g = \gamma_{g-1} \frac{t_{g-1}}{\alpha}$. Note that by definition of α ,

$$\begin{aligned} \gamma_K &= \gamma_1 \prod_{g=1}^{K-1} \left(\frac{t_g}{\alpha} \right) = \gamma_1 \left(\frac{\prod_{g=1}^{K-1} t_g}{\alpha^{K-1}} \right) = \gamma_1 \left(\frac{1/t_K \prod_{g=1}^K t_g}{1/\alpha \alpha^K} \right) = \gamma_1 \left(\frac{1/t_K}{1/\alpha} \right) = \gamma_1 \frac{\alpha}{t_K} \\ &\implies \gamma_1 = \gamma_K \frac{t_K}{\alpha}. \end{aligned}$$

Now choose γ_1 small enough that each $\gamma_g \leq 1$; this holds if

$$\max_{g \in \langle 2, \dots, K \rangle} \gamma_g = \max_{g \in \langle 2, \dots, K \rangle} \gamma_1 \prod_{j=2}^g \left(\frac{t_{j-1}}{\alpha} \right) \leq 1.$$

Denote by $\gamma = (\gamma_1, \dots, \gamma_K)$ the K -tuple just built. Note that $\alpha\gamma$ is a permutation of $\gamma * \mathbf{t}$.

Now by definition of X and Y , by SI and IND, and by SDP

$$\begin{aligned} X \sim Y &\Leftrightarrow \mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t})\overline{X} \sim \alpha \underline{X} \uplus (1 - \alpha)\overline{X} \\ &\Leftrightarrow \gamma * (\mathbf{t} * \underline{X} \uplus (\mathbf{1} - \mathbf{t})\overline{X}) \uplus (\mathbf{1} - \gamma)\overline{X} \sim \gamma * (\alpha \underline{X} \uplus (1 - \alpha)\overline{X}) \uplus (\mathbf{1} - \gamma)\overline{X} \\ &\Leftrightarrow (\gamma * \mathbf{t}) * \underline{X} \uplus (\mathbf{1} - \gamma * \mathbf{t})\overline{X} \sim (\alpha\gamma) * \underline{X} \uplus (\mathbf{1} - \alpha\gamma)\overline{X}. \end{aligned}$$

But the last two districts are equally segregated because $\alpha\gamma$ is a permutation of $\gamma * \mathbf{t}$ and \succsim satisfies GS. Q.E.D.

Proposition 9 *Let $\mathbf{t}^1, \mathbf{t}^2 \in [0, 1]^K$ and let $X = \langle \mathbf{t}^1, \mathbf{t}^2, (1 - t_1^1 - t_1^2, 0, \dots, 0), (0, 1 - t_2^1 - t_2^2, 0, \dots, 0), \dots, (0, \dots, 0, 1 - t_K^1 - t_K^2) \rangle$ be a district. Then there is $\alpha_X \in [0, 1]$ such that $X \sim X(\alpha_X)$. Further, α_X is $\left(\prod_{g=1}^K t_g^1 \right)^{1/K} + \left(\prod_{g=1}^K t_g^2 \right)^{1/K}$.¹⁶*

Proof. The proof is almost identical to the proof of Proposition 6. The only difference is that here the weights are $w_g = 1/K$, and instead of relying on Proposition 5 one needs to rely on the analogous Proposition 8. Q.E.D.

Proposition 10 *For every district X there is a unique $\alpha_X \in [0, 1]$ such that $X \sim \alpha_X \underline{X} \uplus (1 - \alpha_X)\overline{X}$. Further, this unique α_X is $\sum_{n \in N(X)} \left(\prod_{g=1}^K t_g^n \right)^{1/K}$.¹⁷*

Proof. The proof is almost identical to the proof of Proposition 9. The only difference is that here the weights are $w_g = 1/K$, and instead of relying on Proposition 5 and 6 one needs to rely on the analogous Propositions 8 and 9. This ends the proof of the theorem. Q.E.D.

¹⁶This is less than or equal to 1 since the geometric average of a set of numbers can be no greater than their arithmetic average: $\left(\prod_{g=1}^K t_g^1 \right)^{1/K} + \left(\prod_{g=1}^K t_g^2 \right)^{1/K} \leq \frac{1}{K} \sum_{g=1}^K t_g^1 + \frac{1}{K} \sum_{g=1}^K t_g^2 = \frac{1}{K} \sum_{g=1}^K (t_g^1 + t_g^2) \leq \frac{1}{K} \sum_{g=1}^K 1 = 1$.

¹⁷By the reasoning given in footnote 16, α_X must lie between zero and one.

A.3 Proof of Proposition 2

GS: The fact that Group Symmetry is independent of the other axioms follows directly from Theorems 1 and 2: for any vector of weights $\mathbf{w} \neq (1/K, \dots, 1/K)$, the asymmetric Atkinson index $A_{\mathbf{w}}$ represents a segregation order that satisfies SI, WSDP, IND, N, and C, but fails GS.

WSDP: To see that WSDP is independent of the other axioms note that since the Atkinson order satisfies GS, SI, IND, C and N, so does the order represented by the index $1 - A$ (defined by $(1 - A)(X) = 1 - A(X)$). It is clear that this order does not satisfy WSDP.

N: The trivial segregation order, which ranks all districts as equally segregated, violates N while satisfying all the other axioms.

IND: Consider the Unweighted Dissimilarity index D^U . It is clear it satisfies N and GS. It satisfies C since it is represented by a continuous function. SI follows from the fact that for any positive scalar α , $t_g^n = \frac{T_g^n}{T_g} = \frac{\alpha T_g^n}{\alpha T_g}$. WSDP holds since $D^U(X) \geq 0$ for all districts X , and $D^U(X) = 0$ if all the schools of X are representative. As for IND, consider the following districts: $X = \langle (2, 4), (2, 0) \rangle$ and $Y = \langle (4, 2), (0, 2) \rangle$. One computes $D^U(X) = D^U(Y) = 1/2$. Consider now the result of annexing to them the one-school district $Z = \langle (4, 0) \rangle$. One can verify that $D^U(X \uplus Z) = 3/4$ while $D^U(Y \uplus Z) = 1/2$. Hence, D^U violates IND.

SI: The Mutual Information index M clearly violates SI. Since the entropy function is symmetric, M satisfies GS. Since M is continuous, it also satisfies C. That it satisfies WSDP follows from the fact that $M(X) \geq 0$ for all districts X , and that $M(X) = 0$ if all the schools of X are representative. For a proof that the mutual information ordering satisfies IND, see Frankel and Volij [10].

C: Let $\mathbf{w} = (w_g)_{g=1}^K$ and $\mathbf{w}' = (w'_g)_{g=1}^K$ be two different vectors of weights that each sum to one. It is easy to verify that $\succ_{\mathbf{w}, \mathbf{w}'}$ satisfies SI, IND, WSDP, and N since $A_{\mathbf{w}}$ and $A_{\mathbf{w}'}$ do. It clearly violates GS since at least one weight vector must be asymmetric. In addition, it violates C. To see why, let X and Y be two districts with different group distributions such that $A_{\mathbf{w}}(X) = A_{\mathbf{w}}(Y) < 1$ and $A_{\mathbf{w}'}(X) < A_{\mathbf{w}'}(Y)$. Let $c \in (0, 1)$ and consider the district $cX \uplus (1 - c)Y$. Let $\gamma_g = \frac{cT_g(X)}{cT_g(X) + (1 - c)T_g(Y)}$ and $\eta_g = 1 - \gamma_g$. Note that a proportion

$t_g^n(X)\gamma_g$ of group- g students of the district $cX \uplus (1-c)Y$ attend school $n \in N(X)$. Likewise, a proportion $t_g^n(Y)\eta_g$ of group- g students of the district $cX \uplus (1-c)Y$ attend school $n \in N(Y)$. Therefore, we can write

$$\begin{aligned}
1 - A_{\mathbf{w}}(cX \uplus (1-c)Y) &= \sum_{n \in N(X)} \prod_{g \in G} (t_g^n(X)\gamma_g)^{w_g} + \sum_{n \in N(Y)} \prod_{g \in G} (t_g^n(Y)\eta_g)^{w_g} \\
&= \sum_{n \in N(X)} \prod_{g \in G} (t_g^n(X))^{w_g} (\gamma_g)^{w_g} + \sum_{n \in N(Y)} \prod_{g \in G} (t_g^n(X))^{w_g} (\eta_g)^{w_g} \\
&= \left(\prod_{g \in G} (\gamma_g)^{w_g} \right) \sum_{n \in N(X)} \prod_{g \in G} (t_g^n(X))^{w_g} + \left(\prod_{g \in G} (\eta_g)^{w_g} \right) \sum_{n \in N(Y)} \prod_{g \in G} (t_g^n(Y))^{w_g} \\
&= (1 - A_{\mathbf{w}}(X)) \prod_{g \in G} (\gamma_g)^{w_g} + (1 - A_{\mathbf{w}}(Y)) \prod_{g \in G} (\eta_g)^{w_g}.
\end{aligned}$$

Since the group distributions of X and Y are not the same, there are groups $g, g' \in G$ with $\gamma_g \neq \gamma_{g'}$. (Otherwise, for all groups g , γ_g equals a constant λ , which implies $\frac{T_g(X)}{T_g(Y)} = \frac{\lambda(1-c)}{c(1-\lambda)}$. Hence, X and Y must have the same group distribution, a contradiction.) Therefore, the geometric average $\prod_{g \in G} (\gamma_g)^{w_g}$ is strictly lower than the corresponding arithmetic average, and the same is true for $\prod_{g \in G} (1 - \gamma_g)^{w_g}$. As a result,

$$1 - A_{\mathbf{w}}(cX \uplus (1-c)Y) < (1 - A_{\mathbf{w}}(X)) \sum_{g \in G} w_g \gamma_g + (1 - A_{\mathbf{w}}(Y)) \sum_{g \in G} w_g \eta_g.$$

(By assumption, $A_{\mathbf{w}}(X)$ and $A_{\mathbf{w}}(Y)$ are strictly less than one.). Since $A_{\mathbf{w}}(X) = A_{\mathbf{w}}(Y)$, and since c was arbitrary chosen from $(0, 1)$, we obtain that $A_{\mathbf{w}}(cX \uplus (1-c)Y) > A_{\mathbf{w}}(Y)$ for all $c \in (0, 1)$. Consequently the set

$$\{c \in [0, 1] : cX \uplus (1-c)Y \succ_{\mathbf{w}, \mathbf{w}'} Y\}$$

equals $[0, 1)$, which is not closed. Q.E.D.

References

- [1] Atkinson, A.B. 1970. "On the Measurement of Inequality." *Journal of Economic Theory* 2: 244-63.
- [2] Boozer, Michael A., Alan B. Krueger, and Shari Wolkin. 1992. "Race and School Quality Since Brown v. Board of Education." *Brookings Papers on Economic Activity: Microeconomics* 1992, pp. 269-338.
- [3] Chakravarty, Satya R., and Jacques Silber. 2007. "A Generalized Index of Employment Segregation." *Mathematical Social Sciences* 53:185-194.
- [4] Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1982. "Achievement and Segregation in Secondary Schools: A Further Look at Public and Private School Differences." *Sociology of Education* 55:162-182.
- [5] Cotter, David A., JoAnn DeFiore, Joan M. Hermesen, Brenda M. Kowalewski, and Reeve Vanneman. 1997. "All Women Benefit: The Macro-Level Effect of Occupational Integration on Gender Earnings Equality." *American Sociological Review* 62:714-734.
- [6] Cutler, David, and Edward Glaeser. 1997. "Are Ghettos Good or Bad?" *Quarterly Journal of Economics* 112: 827- 872.
- [7] Cutler, David, and Edward Glaeser and Jacob Vigdor. 1999. "The Rise and Decline of the American Ghetto" *Journal of Political Economy* 107: 455- 506.
- [8] Duncan, Otis D., and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indices." *American Sociological Review* 20:210-217.
- [9] Echenique, Federico, and Roland G. Fryer, Jr. 2005. "On the Measurement of Segregation." Mimeo.
- [10] Frankel, David, and Oscar Volij, 2007. "Measuring Segregation." Mimeo.

- [11] Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." Mimeo.
- [12] Hutchens, Robert. 2001. "Numerical Measures of Segregation: Desirable Properties and Their Implications." *Mathematical Social Sciences* 42:13–29.
- [13] Hutchens, Robert. 2003. "One Measure of Segregation." *International Economic Review* 45:555–578.
- [14] Jahn, Julius, and Schmidt, Calvin F., and Schrag, Clarence. 1947. "The Measurement of Ecological Segregation." *American Sociological Review* 12:293–303.
- [15] James, David R., and Karl E. Taeuber. 1985. "Measures of Segregation." *Sociological Methodology* 15:1–32.
- [16] Johnson, Ron, Michael Poulsen, and James Forrest. 2007. Ethnic and Racial Segregation in U.S. Metropolitan Areas, 1980-2000: The Dimensions of Segregation Revisited. *Urban Affairs Review* 42:479-504.
- [17] Lewis, Gregory B. 1996. "Gender Integration of Occupations in the Federal Civil Service: Extent and Effects on Male-Female Earnings." *Industrial and Labor Relations Review* 49:472–483.
- [18] Macpherson, David A., and Barry T. Hirsh, 1995. "Wages and Gender Composition: Why do Women's jobs Pay Less? " *Journal of Labor Economics* 13:426–471.
- [19] Massey, Douglas S. and Nancy A. Denton. 1988. "The Dimensions of Racial Segregation." *Social Forces* 67:281–315.
- [20] Philipson, Tomas. 1993. "Social Welfare and Measurement of Segregation." *Journal of Economic Theory* 60:322-334.

- [21] Reardon, Sean F., and Glenn Firebaugh. 2002. "Measures of Multigroup Segregation." *Sociological Methodology* 32:33-67
- [22] Shorrocks, A. F. 1980. "The Class of Additively Decomposable Inequality Measures." *Econometrica* 39:613-25.
- [23] Shorrocks, A. F. 1984. "Inequality Decompositions by Population Subgroups." *Econometrica* 48:1369-86.
- [24] Taeuber, Karl E., and David R. James. 1982. "Racial Segregation Among Public and Private Schools." *Sociology of Education* 55:133-143.
- [25] Zoloth, Barbara S. 1976. "Alternative Measures of School Segregation." *Land Economics* 52:278-298.